

**ТУКМАКОВА НАДЕЖДА ПЕТРОВНА**

**ОПЫТ СОПОСТАВИТЕЛЬНОГО  
ЛИНГВОСТАТИСТИЧЕСКОГО АНАЛИЗА  
РУССКОГО И ЧУВАШСКОГО ЯЗЫКОВ  
(НА МАТЕРИАЛЕ ГАЗЕТНОЙ ПУБЛИЦИСТИКИ)**

10.02.20 - Сравнительно-историческое, типологическое и  
сопоставительное языкознание

**АВТОРЕФЕРАТ**  
**диссертации на соискание ученой степени**  
**кандидата филологических наук**

Чебоксары 2003

Работа выполнена на кафедре литературного редактирования и стилистики Чувашского государственного университета им. И.Н. Ульянова.

Научный руководитель:

доктор филологических наук  
профессор Петров Н.П.

Официальные оппоненты:

доктор филологических наук  
профессор Сергеев В.И.

кандидат технических наук  
профессор Желтое В.П.

Ведущая организация:

Чувашский государственный педагогический университет им. И.Я. Яковлева

Защита состоится 21 -марта 2003 года в 14:00 часов на заседании диссертационного совета Д 212.301.03. по присуждению ученой степени доктора филологических наук при Чувашском государственном университете им. И.Н. Ульянова: 428034 г. Чебоксары, ул. Университетская, 38.

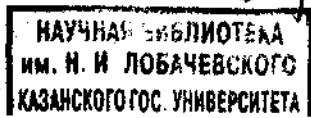
С диссертацией можно ознакомиться в Научной библиотеке Чувашского государственного университета им. И.Н. Ульянова: г. Чебоксары, ул. Университетская, 38.

Автореферат разослан "20" февраля 2003 г.

Ученый секретарь  
диссертационного совета  
доктор филологических наук



А.Р. Губанов



Актуальность исследования. Применение количественного метода в такой, казалось бы, исключительно "качественной" науке, какой является лингвистика, перестало быть спорадическим. Это обусловлено внутренней потребностью языкознания наиболее точно и адекватно изучить свой предмет. Использование в лингвистике математических методов позволяет проникнуть в тайны построения и функционирования языка, а также способствует познанию языка на более высоком уровне и получению объективных знаний о нем. Представляется закономерным обращение к количественному подходу в изучении языковых единиц таких видных ученых, как И.А. Бодуэн де Куртене, А.М. Пешковский, М.Н. Петерсон, Е.Д. Поливанов, В.В. Виноградов.

Своеобразие и особенности языка наиболее отчетливо выражаются в сравнении или сопоставлении его с другими языками. При введении статистического материала результаты такого рода исследования безусловно окажутся более объективными. Одной из задач лингвистической статистики является создание формализованного аппарата, позволяющего достаточно строго описывать типологические различия между языками по данным речи. Такой аппарат должен служить для выявления суммы сходных, и различительных черт, характеризующих системы сопоставляемых языков.

Настоящая работа посвящена сопоставительному изучению русских и чувашских публицистических текстов с использованием количественного метода. Выбор указанных языков не случаен. Общеизвестно: русский язык характеризуется лексико-фразеологическим богатством, разветвленной системой синтаксических структур, обширными стилистическими возможностями. Это язык межнационального общения, оказывающий значительное влияние на контактирующие с ним, системы. Он является в достаточной мере изученным в количественном аспекте и имеет хорошо наработанную литературу. В интенсивных и продолжительных контактах с русским - флективным языком - находится чувашский, агглютинативный по своей природе<sup>2</sup>, также имеющий иные социолингвистические свойства, как правило, отличные от функциональных особенностей русского языка.

Работы сопоставительного характера являются мерилом достаточной изученности языка. Нельзя сказать, что чувашский, как самодостаточная сущность, к настоящему времени изучен абсолютно. Однако первые исследования сравнительно-сопоставительного характера в языкознании появляются уже в середине 20 в. (работы М.Р. Федотова, Н.А. Резюкова). Большую научную ценность в постановке проблем сопоставительного плана между русским и чувашским языками (а именно эти языки являются традиционными компонентами такого рода исследований, практикуемых в

<sup>1</sup> Бсктаев КБ. Статистико-цифровая типология тюркского текста. - Алма-Ата: Наука. 1978. - С. 7-8.

<sup>2</sup> Русский язык получил широкое распространение в среде носителей чувашской культуры. Согласно данным переписи 1989 г., родным пригнали русский 22,3%, чувашей, а 65,4% свободно владеет им.

чувашском языкознании) представляют исследования Н.И. Ашмарина, В.Г. Егорова, И.А. Андреева, Г.Е. Корнилова, М.М. Михайлова, Н.П. Петрова, В.И. Сергеева, М.И. Скворцова, М.Р. Федотова, М.Ф. Чернова, а также современные изыскания З.В. Гриншайновой и А.Р. Губанова. Частные вопросы сопоставительного изучения указанных языков, чаще всего в педагогическом аспекте, отражены в работах А.П. Алексеевой, Г.А. Анисимова, З.Ф. Мышкина, В.И. Орловой, А.С. Уськина. В то же время сопоставительный аспект в чувашском языкознании, несмотря на более чем свою, полувековую историю, находится в состоянии, которое требует уточняющего определения. - "фрагментарный"¹.

В настоящее время средства массовой информации занимают достаточно большое место в жизни человека. Важнейшей их функцией является формирование общественного сознания. Особым доверием пользуются периодические издания как в информативном, так и в стилистико-нормативном плане. Газетная речь, откладываясь в нашем подсознании, является основным источником накопления словарного запаса. Для определения нормы, выяснения степени употребительности и стилистической характеристики лексических, морфологических и синтаксических единиц исследователи все чаще обращаются к газетной речи. Для объективной интерпретации функциональных стилей и во избежание субъективных оценок исследователя газетной публицистики необходимо получить объективные данные с учетом того, что к объектам языка будут применены вероятностные оценки, а сам язык будет интерпретироваться как производящая система с вероятностными ограничениями.

Вследствие ряда причин - малочисленности научного потенциала, нетрадиционности количественного направления - чувашская газетная речь еще не описана в статистическом плане, не говоря уже о сравнительно-сопоставительном аспекте. Между тем чувашская публицистическая речь сформировалась под влиянием русской газетной публицистики. Большинство информационных материалов редакции средств массовой информации получают на русском языке или непосредственно из русских источников. Нередко наблюдается профессиональное чувашско-русское двуязычие. Наблюдение над феноменом взаимодействия названных языков может привести к интереснейшим выводам.

Целью исследования настоящей работы является создание объективной картины современного состояния русской и чувашской газетной речи; нахождение точек соприкосновения и линий расхождения на морфологическом, синтаксическом и лексическом уровнях между языковыми единицами русских и чувашских газетных текстов.

Для достижения указанной цели необходимо решить следующие задачи:

1) определить однородность и объем репрезентативных выборок русских и чувашских газетных текстов;

---

¹ Ведение компаративных разысканий с равным привлечением чувашского материала было и, скорее всего, останется прерогативой чувашских ученых.

2) установить вероятности употребления каждой части речи, выявить корреляционную связь между ними, а также создать вероятностную модель взаимодействия частей речи в русских и чувашских газетных текстах;

3) определить среднюю длину предложения и установить связь между величиной предложения и его структурой;

4) установить коэффициент диверсификации, а также составить минисписок наиболее частотных слов в изучаемых газетах.

**Объект изучения** - публицистическая разновидность употребления русского и чувашского языков.

**Предмет исследования** - количественные закономерности языковых единиц лексического и грамматического уровней, участвующих в организации публицистического стиля.

Объект исследования и поставленные задачи предопределили применение следующих **методов**:

1) статистического - для установления количественного выражения языковых единиц;

2) синхронного сопоставительного анализа - для выявления общих и отличительных явлений между сравниваемыми языками в их публицистической выраженности;

3) стилистического - с целью установления адекватности выражения оттенком основного значения слова, для определения жанровых разграничений русских и чувашских газетных текстов;

4) индуктивного заключения - для обобщения результатов проделанного эксперимента.

**Методологическую** базу диссертационного сочинения составили работы Н.Д. Андреева, А.Н. Бектаева, Б.Н. Головина, М.Н. Кожиной, А.Н. Колмогорова, А.А. Маркова, А.И. Носенко, Р.Г. Пиотровского, Р.М. Фрумкиной и др.

**Источниками** иллюстративных примеров послужили тексты республиканских газет "Советская Чувашия" ("СЧ") и "Хыпар" ("Х"), ограниченные хронологическими рамками с 1998 по 2002 гг. общим объемом в 42 тыс. словоупотреблений. Выводы основываются на изучении стилоопределяющих жанров публицистической разновидности употребления языка: информационных (ИЖ) - событийная заметка, информационная корреспонденция, информационный отчет; аналитических (АЖ) - статья, аналитическая корреспонденция, аналитический отчет.

**Научная новизна** работы определяется тем, что в ней впервые исследуются русские и чувашские газетные тексты в сопоставительно-статистическом аспекте, определяются такие параметры сравниваемых газетных текстов, как предметность, качественность, динамизм, корреляция, сложность и лексическое богатство. В диссертации уточняются и развиваются имеющиеся в трудах чувашских лингвистов (Н.П. Петрова, А.П. Хузангай, В.В. Андреева, В.И. Котлеева, В.П. Зайцевой) качественно-количественные исследования чувашского языка и его функциональных разновидностей..

**Теоретическая значимость исследования** состоит в том, что такого рода исследования способствуют введению в количественную лингвистику чувашского материала, до последнего времени игнорируемого в общетеоретических изысканиях как следствие недостаточности его количественного описания. Возможно обращение к статистическому материалу, изложенному в настоящей работе, при построении вероятностной модели русской и чувашской публицистической речи, создание которой необходимо для дальнейшего развития теории журналистики и совершенствования газетной практики.

**Практическая значимость** исследования заключается в том, что результаты работы могут быть использованы в работе переводческой службы в средствах массовой информации Республики, а также при подготовке спецкурса и составлении учебного пособия по сопоставительному изучению русских и чувашских языков на филологических специальностях высших учебных заведений. Исследование имеет большое практическое значение при изучении языка, не являющегося родным (на продвинутом этапе).

#### **На защиту выносятся:**

- 1) способы выражения категории количества и ее место в установлении вероятностной модели естественного языка;
- 2) результаты сопоставительного лингвостатистического анализа русской и чувашской газетной речи на грамматическом уровне;
- 3) величина коэффициента диверсификации в исследуемых региональных газетах;
- 4) выявление общих и отличительных черт между русской и чувашской газетной публицистикой.

**Апробация работы.** Основные положения диссертационного исследования докладывались на ежегодных научных конференциях Чувашского государственного университета им. И. Н. Ульянова, посвященных 60-летию со дня рождения доктора филологических наук, профессора В.И. Сергеева и 30-летию со дня образования университета.

По теме диссертации имеется 5 публикаций в сборниках научных трудов.

**Структура и объем работы.** Диссертация состоит из введения, двух глав, заключения, списка использованной литературы, примечания и приложений.

## **СОДЕРЖАНИЕ РАБОТЫ**

Во **введении** дается представление о проблематике работы, определяются ее актуальность, новизна, цели, задачи, теоретическая и практическая значимость, подчеркивается целесообразность исследования проблем чувашского языка в сопоставлении с русским языком, а также признается необходимым обращение к статистическим методам для устранения вкусового субъективизма.

В первой главе "**Место категории количества в установлении вероятностной модели естественного языка**" обосновываются реальные,

существующие в любом естественном языке основания для применения статистических методов; приводятся примеры имплицитной и эксплицитной количественной выраженности языковых единиц в русском и чувашском языках; прослеживается взаимосвязь основных стилистических систем с категорией количества.

Качество и количество являются универсальными, взаимообусловленными категориями. Качеством называется то, что образует своеобразие какого-то материального объекта, а в лингвистике - языкового факта. Качественная определенность не представляет собой нечто готовое: она постоянно развивается. По мере ее развития обособляется тот аспект, который называется количеством. "Число оказывается ничем иным, как отношением множеств"<sup>1</sup>. С этой точки зрения оно оказывается не просто математическим, но и философским понятием. Любой естественный язык является субстанцией, а потому имеет внешнюю форму проявления материальности в виде букв, т.е. обладает качественной определенностью - разнообразием, множеством. Количественная определенность выражена в парадигматических и синтагматических отношениях между единицами языка.

Парадигматическая система языковых единиц, на первый взгляд, составляет исключительно царство качественности. Но, по замечанию В.Г. Адмони, "каждое языковое явление" помимо качественной определенности, обладает и некоей "массой", т.е. занимает в кругу других языковых явлений некое пространство и соизмеримо с ними в неких измерениях с разной степенью интенсивности<sup>2</sup>. Проанализируем отношения между синонимами в синонимическом ряду в чувашском и русском языках: *чуп (бежать) - варкан (устремляться, стремительно нестись) - ыткан (бросаться, кидаться, мчаться) - веҫ (лететь)*. Количество в них выражено через отношение синонимов друг к другу. Степень интенсивности признака в каждом последующем слове выражена больше, чем в предыдущем. Если представить отношения между словами математически, то получится:  $a < b < c$ .

Разные по значению слова также могут отражать различную интенсивность какого-либо определенного свойства. Яркие примеры демонстрируют прилагательные с различной степенью освещенности: в чувашском языке - *чакар (светло-голубой), сенкер (голубой), сеп-сенкер (чисто голубой), кавак (синий, голубой, серый, фиолетовый), кан-кавак - (синий-синий)*; в русском языке - *голубой, голубоватый, темно-голубой, синий, синеватый, темно-синий, фиолетовый*. Указанная система слов обозначает некий общий признак, встречающийся с различной интенсивностью (от "голубой" до "фиолетовый").

Сравнительная и превосходная степени прилагательных и наречий обозначают качество предмета в сравнении. В этом отношении они связаны с

<sup>1</sup> Ахлибинский В.В., Храленко Н.И. Теория качества в науке и практике.- Л.: Изд-во Лeningr. ун-та. 1989.- С. 120.

<sup>2</sup> Адмони В.Г. Качественный и количественный анализ грамматических явлений // Вопр. языкознания.- 1963.- №4.- С. 57.

мыслью, отражающей количество. Качественные прилагательные и в русском, и в чувашском языках имеют три степени сравнения: положительную, сравнительную и превосходную. Положительная степень не предполагает сравнения качества в предметах, поэтому не выражает количества. Ср.: *шура(белый)* - *шурарах (белее)* - сравнительная степень; *шур (белый)* - *шура—шура, шап—шура(белый—пребелый)*; *чи шура(самый белый)*; *тёлёммелле шура(удивительно белый)*; *калама сук шура (невыразимо белый)*; *пуринчен шур (белее всех)* - превосходная степень.

В чувашском и русском языках все наречия, образованные от качественных прилагательных, также имеют сравнительную степень: *х в рт - хавартрах (быстро ~ быстрее)*; *шап - шапрах (тихо - тише)*. В чувашском языке наречия, имеющие сравнительную степень, могут быть и в превосходной: *хаварт - хап-хаварт (очень быстро)*; *шап - шап-шапарт (очень тихо)*.

К сравнению (а значит, и отражению мысли о количестве) относятся также увеличительные и уменьшительные суффиксы. В русском языке при помощи специальных суффиксов к основному значению слова может добавляться значение сравнительной величины предмета: *заяц - зайчик - зайчишка — зайныка - зайчонок*; *изба - избушка - избенка*; *дом - домик - домишко - домище*. В чувашском языке существительные не имеют увеличительных и уменьшительно-ласкательных аффиксов. Для обозначения "количества" существительного используются слова *пёчэк (маленький)*, *пыс к(большой)*, *в там (средний)*: *пёчэк мулкач, мулкач сури - зайчонок, зайчик, зайчишка*; *пёчэк сурт - маленький дом, домик*; *пыс к сурт - большой дом, домище*.

Аффиксы, имплицитно выражающие количество (или интенсивность), в чувашском языке могут иметь только прилагательные. Таковыми являются аффиксы —шка(-шке), образующие имена прилагательные от прилагательных: *к тра (кудрявый) - к трашка (кучерявый, волнистый)*; *семсе (мягкий) — семсешке (слабенький, хиленький)*; *лутра (низкорослый) — лутрашка (низенький)*. Таким образом, сфера парадигматических отношений языковых единиц содержит в себе как качественные, так и количественные моменты.

Если перейти к другой сфере языковых единиц - сфере употребления, сцепления, сочетания и организации их в процессе речевого функционирования, то здесь легко обнаружатся как качественные, так и количественные моменты. Язык находится в употреблении, и частотность лингвистических объектов различна в зависимости от ситуации, в которой происходит непрерывный отбор единиц и категорий языка. Количественность в синтагматике языка имеет более явную выраженность, чем в парадигматике, и обладает свойством стилиобразующей активности, т.е. способностью различать функциональные стили языка.

В языкознании существует множество определений понятия "стиль". Однако в каждом из них удавливается общий признак — специфическая форма сообщения, "устная или письменная. Всякое отклонение может быть



измерено и подвергнуто математическому анализу. В стилистике отклонение выражается в наличии в каждом из стилей специфических языковых средств. Этот признак достигает максимума в индивидуальном стиле, который есть мера отклонения от нейтральной нормы. Понятие стиля, таким образом, латентно содержит в себе понятие количества.

Лингвистический знак, в противоположность математическому, наряду с основным значением - денотацией, характеризуется дополнительными оттенками - коннотацией, но к нему также (а по Ш. Балли - именно поэтому), как и к знаку математическому, применимо понятие функции. Ш. Балли показывает, что коннотация и денотация связаны с двойной функцией языка: объективной, понятийно-логической, с одной стороны, и субъективной, аффективной, - с другой, поэтому истоки стилистических эффектов следует искать в системе языка.

Язык не может существовать вне функционирования, как не может существовать материя без движения. Функция является одним из основных понятий математики, выражающее зависимость одних переменных величин от других. Считая возможным приложить математическое понятие функции к лингвистике, М.В. Мачавариани исходит из того, что "так как язык есть система, то отсюда а priori следует вывод, что отдельные элементы этой системы взаимосвязаны; отсюда возможная эффективность применения понятия *функции*". Речь является функцией языка. Язык и речь - реальные объекты, взаимно предполагающие друг друга и взаимодействующие в речевой деятельности. Качественные изменения в любом языке происходят путем постепенного учащения или, наоборот, сокращения употребительности соответствующих языковых элементов в стилях речи. Язык предоставляет средства, а речь осуществляет их выборочный набор в зависимости от ситуации, то есть от экстралингвистических факторов: тематики общения, ведущих актуальных коммуникативных задач, целевой установки и т.д.

В настоящее время во всех языках центральное положение занимает обиходно-литературный стиль, т.к. бытует в повседневном общении. Во многих современных литературных языках выделяются публицистический, производственно-технический, официально-деловой и научный функциональные стили. Все эти стили дифференцируются на фоне обиходно-литературного стиля. Основная масса грамматических явлений может свободно употребляться в любом из них, т.е. носит нейтральный характер по отношению к стилям, однако степень активности одних и тех же элементов языка в разных стилях различна. По-разному функционируют языковые явления и в различных языках, т.к. у каждого народа свои

---

<sup>4</sup> Про П. Разделы и направления стилистики и их проблематика // Новое в зарубежной лингвистике: Лингвистилистика. М.: 1979. - Вып. 9. - С. 41.

Мачавариани М.В. О взаимоотношении математики и лингвистики. Всп. языкознания. - 1963 - № 3 - С. 83-91.

особенности мышления<sup>1</sup>. Закономерности функционально-стилистической системы, таким образом, носят вероятностно-статистический характер, а мысль об использовании количественных методов для характеристики функциональных стилей "лежала на поверхности" и обусловлена внутренней потребностью языкознания наиболее точно и адекватно изучить свой предмет.

Понимание функциональных стилей только как отклонения от некоей нормы, безусловно, не является полным. Стил не может быть отвлеченным понятием: он всегда существует в каком-нибудь конкретном тексте. Следовательно, нельзя акцентировать внимание только на стилистически маркированных элементах. Необходимо помнить и об элементах стилистически нейтральных, которые, как правило, преобладают во всяком тексте. Кроме того, стил не может быть просто совокупностью каких-либо элементов, а представляет собой организованную последовательность, систему сознательно отобранных элементов. Наконец, стил в какой-то мере идентичен с понятием количества в том смысле, что способен быть измеряемым. "Об измерении можно говорить только тогда, когда исследователю удастся разумно атомизировать это сложное явление и каждому элементарному значению признака приписать соответствующую величину"<sup>2</sup>. Соответствующая величина в лингвистике, в зависимости от целей исследования, может быть представлена в результате количественного подхода к языковым единицам на самых разных уровнях любого языка. Такие квантитативные исследования, как вычисление отношений числа гласных к числу согласных, определение среднего числа фонем, приходящихся на слог, слово, количество словообразовательных гнезд и т.д., не могут считаться важными для объективной характеристики функциональных стилей. Стилистические особенности языка проявляются на более высоких лингвистических уровнях - лексическом, морфологическом и синтаксическом. Понятие стили - явление речевое, поэтому стил надо считать свойством речи, текста, а не языка. Язык обладает лишь стилистическим потенциалом, разными возможностями стилиобразования. Текст соотносится с языком через функциональный стил. Под вероятностной природой функциональных стилей прежде всего подразумевается вероятностный характер текстов, т.к. именно в текстах количественная определенность языковых единиц представлена как данность. Текст измеряем благодаря своей фиксации. Он представляет собой легко обозримую систему использования языковых единиц. Стил

---

<sup>1</sup> В русском языке именительный и родительный падежи встречаются в несколько раз чаще творительного и во много раз чаще дательного. В чувашском же языке самым употребляемым является основной падеж имен существительных, т.к. он способен выполнять функции разных падежей. В газетных текстах, как показывают статистические исследования В.П. Зайцевой, наиболее распространенной формой является дательный падеж [Андреев Н.Д., Зинлер Л.Р. О понятиях речевого акта, речи, речевой вероятности и языка // Вопр. языкознания. - 1963. - № 3. - С. 18-19; Зайцева В.П. Исследования частотности употребления слов в различных типах чувашских текстов: Дие. ... канд. филол. наук: 10.02.06. - Чебоксары: Чуваш. гос. Ун-Т, 2000. - С. 112].

<sup>2</sup> Мартыненко Г.Я. Основы стилиметрии. - Л.: Изд-во Лeningr. ун-та. 1988. - С. 10-11.

выражается в тексте как общее через частное, а текст - как единичное в общем. Между понятиями язык и речь существует некая норма, как некий эталон статистического построения речи, изначально заложенный в носителе языка.

Понятие жанра в статистическом аспекте также оказывается существенным, т.к. является одним из компонентов иерархии стилистических систем и основных текстологических единиц: функциональных, жанровых стилей и типов текста. Соотношение между ними - это последовательное подчинение, ибо каждый функциональный стиль охватывает несколько жанровых стилей, каждый из которых включает в себя набор стилей типов текста<sup>1</sup>. Жанр является стимулятором выбора определенных языковых единиц в тексте, а также способствует их определенной частотности. Абстрактное понятие "функциональный стиль", проявляясь в конкретных текстах, фиксирует как нормы литературного языка, так и стилистический отбор, продиктованный ему жанром.

Измерение текста способствует в какой-то мере установлению вероятностных моделей языка. Применение количественного метода к текстам открывает путь к изучению самого языка, так как сегменты текстов, являющиеся объектами подсчетов, относятся к единицам языка, а также способствует выявлению качества более высшего порядка. Таким образом, значение категории количества для лингвистики обнаруживается через отношение текста к языку.

Язык представляет собой систему сложную и неоднозначную. Изучить такую систему целиком - значит объять необъятное. Любое лингвистическое описание языка неизбежно является схематичным, приблизительным, с большей или меньшей вероятностью того, что он (язык) исследован совершенно. Язык так или иначе соприкасается с понятием вероятности, отличительной чертой которой является великодушие, выражающееся в том, что оно предоставляет свои методы и приемы для решения многих проблем во многих областях различной природы; Следовательно, оно содержит в себе и философский аспект, чем и объясняется его универсальность. Вообще, вероятность представляется тождественной философским понятиям абсолютной и относительной истины. Абсолютная истина - знание всего языка - человеку недоступна: возможно лишь постепенное приближение к ней через относительные истины, а именно через реализацию языковых единиц в процессе коммуникации. От теории вероятностей лингвистика воспользовалась двумя важными видами услуг: во-первых, через обнаружение закономерностей в количественном анализе языковых единиц способность предугадывать их поведение в будущем, что является важным для лингвистики моментом вследствие безудержного развития языка; во-вторых, возможность строить общие выводы о языке на основе частного проявления языковых единиц. По отношению к языку, как к системе, текст

---

<sup>1</sup> Ризель Э.Г. К вопросу об иерархии стилистических систем и основных текстологических единиц // Иностр.яз.вук.- 1975.-№6.-С.8. 11.

является его аналогом, вероятностной моделью. Последний факт имеет решающее значение для лингвостатистиков: модель должна максимально точно отобразить свой объект. Задача лингвиста - как можно ближе приблизить вероятностный аналог (текст) к натурному объекту (языку) при невозможности изучить этот объект абсолютно. Статистические данные языка изначально носят вероятностный, случайный характер; Поэтому при интерпретации количественных данных лингвисту часто не обойтись без модальных слов "может быть, должно быть, наверное, видимо, вероятно, по всей вероятности, видно, скорее всего" и т.п. Этот факт дает основание игнорировать применение количественных методов для анализа языковых единиц. Однако данный метод обладает некоторым преимуществом перед другими способами описания языка в том смысле, что в статистически исследованных случайных явлениях обнаруживается некоторое объективно существующее свойство, которое имеет тенденцию оставаться постоянным и проявляться все яснее при увеличении объема исследуемого материала. Указанное свойство измеряется некоторой постоянной величиной, которая является количественной объективной числовой характеристикой изучаемого явления<sup>1</sup>.

Во второй главе "Лингвостатистический анализ языковых единиц русских и чувашских текстов" описывается процесс проведения количественных измерений текста на морфологическом, синтаксическом и лексическом уровнях и производится интерпретация полученных данных. Полученные результаты выявили черты сходства и различия между русской и чувашской публицистикой, а также типологические особенности языков..

В ИЖ газет обнаружилось преобладание имен существительных, что свидетельствует об именном строе данного жанра, о принадлежности его к речи рассуждающей. Главное место в нем занимают задачи разъяснения и называния. Высокочастотными частями речи как в "СЧ", так и в "Х" явились существительные, прилагательные и глагольные формы. Удельный вес прилагательных связан, скорее всего, с большой ролью существительных и, в свою очередь, характеризует книжный характер жанра. Количество глагольных форм в чувашских текстах значительно превышает их количество в русских: на 278 глагольных форм в "СЧ" приходится 420 в "Х".

АЖ является повествовательным, описательным. По логике, он должен иметь вербальный характер. Однако существительные по-прежнему преобладают над остальными частями речи. Прилагательные же в данном жанре стали играть меньшую роль: их количество значительно поубавилось. Но стало больше глагольных форм. При этом их число в "Х" выше более чем в два раза в сравнении с "СЧ". В создании газетного "микрообраза" в обеих газетах важную роль играют существительные прилагательные и глагольные формы. Разница состоит только в их соотношении между собой. Составим

---

<sup>1</sup> Цотровский Р.Г. и др. Математическая лингвистика, - М.: Высш. шк., 1977 - С. 120.

статистическую картину исследуемых газет с помощью вычисления таких понятий, как предметность, качественность и динамизм.

Предметность может измеряться отношением числа имен существительных и местоимений к числу прилагательных и глаголов:

$$П = \frac{Сущ + Мест}{Прил + Гл}$$

Качественность вычисляется отношением числа прилагательных и наречий к числу существительных и глаголов:

$$К = \frac{Прил + Нар}{Сущ + ГЛ}$$

Динамизм допустимо определить через отношение глагольных форм к существительным, прилагательным и местоимениям:

$$Д = \frac{Гл + Прич + Деесп.}{Сущ + Прил + Мест}$$

Результаты вычислений см. в таблице 2.1.

Таблица 2.1.

Количественная выраженность предметности, качественности и динамизма в русских и чувашских газетных текстах

	Информационный жанр		Аналитический жанр	
	"Советская Чувашия"	"Хыпар"	"Советская Чувашия"	"Хыпар"
Предметность	1.63	3.21	1.74	2.5
Динамизм	0.24	0.34	0.24	0.47
Качественность	0.38	0.22	0.36	0.22

Региональная русская публицистика характеризуется, в первую очередь, предметностью, затем качественностью и, наконец, динамизмом. Чувашской публицистике свойственно иное расположение: предметность, динамизм, качественность. Чувашские глагольные формы обладают большей морфологической и эмоциональной нагрузкой, темп чувашской газетной речи, в отличие от русской, более динамичен, действенен<sup>2</sup>.

Простой подсчет, связанный со словоклассификацией, дает неполное представление о функционировании частей речи в текстах различных стилей и жанров. Гораздо полезнее установить степень статистической спаянности между ними. Для большей достоверности результатов было применено два способа для ее установления - коэффициент ранговой корреляции, который находится по формуле Спирмена:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

<sup>1</sup> Головин Б.Н. Язык И статистика. - М.: Просвещение, 1971. - С. 147-148.

<sup>2</sup> Глагольность чувашской речи осознавалась учеными давно, но носила интуитивный характер, не доходя до эмпирического осмысления.

и коэффициент взаимной сопряженности, имеющий следующий вид:

$$K = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$$

Результаты вычислений представлены в таблице 2.2:

Таблица 2.2.

Величины коэффициентов ранговой корреляции между частями речи  
в русских и чувашских газетных текстах<sup>2</sup>

Коррелируемые части речи	Информационный жанр		Аналитический жанр	
	"Советская Чувашия"	"Хыпар"	"Советская Чувашия"	"Хыпар"
Сущ. - гл.	- 0.39/- 0.41	-	-	- 0.49/- 0.5
Сущ. - прич.	+ 0.38/+ 0.41	+ 0.42/+ 0.2	-	+ 0.47/+ 0.22
Сущ. - прил;	+ 0.76/+ 0.82	+ 0.22/+ 0.53	+ 0.42/+ 0.41	+ 0.22/+ 0.35
Сущ. - предлог / послелог	+ 0.64/+ 0.41	-	+ 0.16/+ 0.41	-
Гл. - прил.	-	+ 0.6/+ 0.75	- 0.37/- 0.58	-
Гл. - дееприч.	-	+ 0.19/+ 0.17	-	+ 0.59/+ 0.4
Прил. - наречие	-	+ 0.81/+ 0.5	-	- 0.82/- 0.78
Прич. - деепр.	-	+ 0.78/+ 0.35	-	+ 0.36/+ 0.27

Сводные данные о величинах ранговой корреляции между частями речи свидетельствуют о том, что:

- между существительными и глаголами в ИЖ русской газеты и АЖ чувашской обнаружилась отрицательная статистическая связь, проявляющаяся в том, что при увеличении частоты имен существительных частота глаголов уменьшается и наоборот: при уменьшении существительных возрастает частота глаголов. В ИЖ "Х" и АЖ "СЧ" обнаружилась другая тенденция - функциональная статистическая независимость указанных частей речи<sup>3</sup>;

— в ИЖ русской и чувашской газет, а также в АЖ газеты "Х" наблюдается прямая статистическая связь между существительными и причастиями. Наличие указанной статистической связи для русской газеты может быть интерпретировано ее ориентированностью на книжную речь: *Между тем только с батыревским хозяйством "Труд" заключен договор на хранение в "Пахча" с правом продажи 600 тонн картофеля и 300 тонн моркови* [Нужен ли столице Большой ОПТ // СЧ.— 2001.- 4 авг.]; *Подведены итоги экономического соревнования между крупными и средними предприятиями (организациями) промышленности республики за первое полугодие. За*

<sup>2</sup> Носенко И.А. Начало статистики для лингвистов. - М.: Высш. шк., 1981. - С 118-119.

<sup>3</sup> При внесении полученных результатов в таблицу соблюдалось два принципа: во-первых, расхождение между величинами коэффициентов ранговой корреляции и взаимной сопряженности должно быть небольшим; во-вторых, сами цифры должны быть значимыми.

<sup>3</sup> Головин Б.Н., исследуя корреляционную связь между существительными и глаголами в текстах А.И.Герцена, И.С.Тургенева и А.И.Гончарова, высказал предположение, что чем независимее друг от друга указанные части речи, тем более речь писателя рационалистичнее и ближе к публицистике [Головин Б.Н. Опыт применения корреляционного анализа в изучении языка // Вопросы статистической лингвистики. - Киев, 1974. - С. 13].

стабильную работу и достижение высоких экономических показателей предприятиям присвоены призовые места [Кто лучший? // Там же.- 2001.- 16 авг.]. Для чувашской газетной речи употребление причастий не является стилистически маркированным, т.к. глагольные формы в чувашском языке имеют большой удельный вес независимо от функциональной принадлежности;

- во всех жанрах газет выявилось тесное функциональное притяжение существительных с прилагательными. Полученные результаты дают положительный ответ на интересующий многих лингвистов вопрос: действительно ли прилагательные являются лишь покорными спутниками существительных? Цифры показывают: активность прилагательных зависит от активности существительных. Такая зависимость приводит к образованию устойчивых сочетаний, что является характерной чертой газетной публицистики. В русских газетах в силу своей высокочастотное™ устойчивыми можно назвать такие сочетания, как *активное участие, основная функция, реальные доходы, высокая награда, торжественная церемония* и т.д. В чувашских газетах - *пысак уяв (большой праздник), пысак зал (большой зал), пысак ёс (большая работа), чапла церемони (торжественная церемония), хакла тавансем (дорогие родственники), ыра ханасем (добрые гости), полномочиллĕ представителе (полномочный представитель), суйлав кампанийĕ (предвыборная кампания)* и т.д.;

- величина степени статистической спаянности между существительными и предложениями в русской газете демонстрирует их положительную связь в обоих жанрах. Это объясняется особенностями языка газеты, характерными для него конструкциями "существительное + существительное с предлогом", такими как *договор на хранение, плата за услуги, тариф за пользование* и др. Для чувашской газетной речи свойственна функциональная независимость послелогов от существительных — конструкция "существительное + существительное": *бюджетна бюджет (бюджет с бюджетом), саккунпа саккун (закон с законом), республик па республика (республика с республикой), патшаллахпа патшаллах (государство с государством)*;

- между глаголами и деепричастиями в обоих жанрах "X" выявилась положительная статистическая связь. Данный факт определяется структурными особенностями чувашского языка, для которого характерно обилие аналитических конструкций: *Президент СССР халăх артисткин В.К. Кузьминан какарĕ сине орден тирсе хучĕ* - Президент прикрепил орден к груди заслуженной артистки СССР В.К. Кузьминой [Клементьев И. Ёсĕре паларнисене чысларĕс // X.- 1999.- 30 дек.]; *Тупаш астан ытларах илесси хăех куранса пырать. Укса аш, сĕр улми, пахча симĕс сутса тавассĕ. Темисе сул каялла кана—ха сурт хысĕнче сĕр улми кана устеретчĕс, халĕ пахча симĕс ытларах туса илессĕ* [Михайлов Ю. Сер пайĕпе ёслеме халлĕхе хатĕр мар // Там же.-2000.-2 авг.] - Откуда получить прибыль - выясняется само по себе. Деньги делают на продаже мяса, картофеля, овощей. Буквально

несколько лет назад за домом выращивали только картошку, ныне получают больше овощей;

- в чувашских текстах имеет место наличие прямой связи между причастиями и деепричастиями: *2001сультта продукция туса к ларасси хаварт чакса ларна.. Укса-тенкё тёлёшёнчен пысак йыварлахсем тухса тана* [Павлова И., Алексеев Л. Сынсемпе - куса-кусан // Там же. - 2002. - 30 апр.] - В 2001 году выпуск продукции резко снизился. С финансовой стороны возникли большие трудности; *Фойерен тытанса чи султи картлашкасен икё ёнёне чаваши тумё таханна хёрупрас тарса тухна* [Григорьев В. Президент халах ырлахёшён ёсleme тупа турё // Там же. - 1998. - 13 янв.] - Начиная с фойе, по обеим сторонам самых высоких ступенек выстроились девушки в чувашских национальных костюмах;

- величины корреляции между прилагательными и наречиями в "Х" получились контрастными - прямая в ИЖ и обратная в АЖ. ИЖ является номинативным, в уточнении существительных важную роль играют прилагательные. В АЖ признак, скорее всего, достигается наличием более тесной статистической спаянности глаголов с деепричастиями. В чувашском языке вспомогательные глаголы придают дополнительный оттенок деепричастию, благодаря чему составные глаголы получают новые семантические качества. Они как бы живописуют действие, характеризуя его наглядно, образно, картинно;

- отличительные черты, связанные со структурными особенностями языков, отразились в употребляемости служебных частей речи - союзов и предлогов (в русском), послелогов (в чувашском). Для русской публицистики характерна прямая статистическая связь предлогов практически со всеми частями речи. В чувашской публицистике она в большинстве случаев близка к нейтральной, что объясняется особенностями этого языка: соединительную функцию между частями речи в нем могут выполнять различные аффиксы. В частотном соотношении количество послелогов в чувашских текстах значительно уступает функционально равноправным русским предлогам: в ИЖ "СЧ" - 245, "Х" - 59, в АЖ "СЧ" - 363, "Х" - 91. В русских текстах союзы употребляются гораздо чаще, чем в чувашских: в ИЖ "СЧ" - 165, АЖ - 250; в ИЖ "Х" - 35, АЖ - 93. В чувашском языке в определенных случаях функцию союза выполняет аффикс творительного падежа -па (- пе) и аффикс деепричастия -са (- се).

Средняя длина предложения отражает манеру изложения. При всей свободе в выборе способа изложения мыслей пишущий исходит от экстралингвистических факторов. Поэтому влияние длины предложения на создание объективной картины функциональных стилей безусловно. Средняя длина предложения измерялась количеством входящих в него слов и находилась путем деления общего количества словоупотреблений в выборке на количество предложений в ней.

Средняя длина предложения "СЧ" в обоих жанрах оказалась больше аналогичного показателя в "Х" (ИЖ "СЧ" - 157.5, "Х" - 128.2; АЖ "СЧ" - 134.8, "Х" - 126.4). Однако обе газеты отдают предпочтение изложению



событий и фактов в простых и простых распространенных предложениях (в ИЖ русской газеты они составили 62%, в АЖ - 65%; в ИЖ чувашской - 79%, в АЖ - 83%). Чтобы убедиться в достоверности полученных величин, по формуле коэффициента линейной корреляции:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(n-1)s_x^2 \cdot s_y^2}} \quad 1$$

вычислим величину корреляционной связи между количеством предложений всех типов и количеством простых и простых распространенных предложений (ИЖ "СЧ" - 0,48, "Х" - 0,95, АЖ "СЧ" - 0,89, "Х" - 0,91), между количеством предложений всех типов и количеством сложных предложений (ИЖ "СЧ" - 0,682, "Х" - 0,21, АЖ "СЧ" - 0,15, "Х" - -0,17)<sup>2</sup>. Результаты свидетельствуют о том, что структура предложения в АЖ обеих газет и ИЖ "Х" влияет на его длину: чем больше в выборках количества предложений всех типов, тем больше в них количества простых и простых распространенных предложений и тем короче сами предложения. В ИЖ "СЧ" наличие статистической связи между количеством предложений всех типов с количеством простых и простых распространенных оказалось недостоверным, с количеством же сложных предложений обнаружило тесную положительную связь. Влияние структуры предложения на его длину в указанном жанре "СЧ" оказалось менее прогнозируемым: характер предложения не всегда влияет на его величину. Полученные цифры отражают как своеобразие языков с различной типологией, так и их жанровые отличия. Русские предложения строятся на основе сочинительных связей, чувашские — подчинительных; Чувашский язык по своей структуре является аналитическим языком, для которого характерны расчлененность высказывания, ослабленность синтаксических связей, сжатие и опрощение синтаксических конструкций. Сходство газет проявилось в жанровых разграничениях: АЖ менее официален в своих изложениях и более близок к разговорной речи.

Коэффициент диверсификации способствует установлению степени лексического разнообразия текста и определяется отношением числа разных слов к числу всех словоупотреблений:  $C = V/N$ .

При сравнении полученных цифр налицо меньшая величина коэффициента диверсификации в чувашских текстах<sup>3</sup>. В АЖ обеих газет эта

<sup>1</sup> Носенко И. Л. Начало статистики для лингвистов. - М.: Высш. шк., 1981. - С. 125.

<sup>2</sup> В данной работе было взято по 10 выборок из каждого жанра. Критическое значение для такого количества выборок по таблице "Критические значения коэффициента корреляции" равно 0,632 (при степени надежности 0,05). Если полученное значение меньше табличного, то корреляционная связь сомнительна [Головин Б. Н. Из курса лекций по лингвистической статистике. - Горький: Изд-во Горьк. гос. ун-та. 1966. - С. 127].

<sup>3</sup> Андреев И. А. считает неправомерными суждения о богатстве или бедности языка по числу представленных в нем словарных единиц, т. к. один и тот же объем информации может быть выражен различным количеством слов, поэтому объем словаря сам по себе может свидетельствовать не о богатстве

величина оказалась меньше по сравнению с ИЖ (ИЖ "СЧ" - 7.02, "Х" - 6.65; АЖ "СЧ" - 6.25, "Х" - 5.43). По логике в ИЖ степень лексического разнообразия текстов должна быть выражена меньшей величиной по сравнению с АЖ. По крайней мере, в ИЖ часто повторяющиеся слова были бы объяснимы: закрепляясь в употреблении, они помогают легче улавливать смысл и экономят время. Задача этого жанра - сжато, доступно и однозначно передать информацию. АЖ является в сущности неофициальным, ему предоставлена большая свобода в выборе языковых средств. Полученные результаты скорее говорят не о лексической бедности АЖ по сравнению с ИЖ, а выражают математическую закономерность: чем больше словоупотреблений, тем больше лексическая покрываемость уже имеющихся слов. Начало текста характеризуется обилием новых слов. К концу текста вероятность их появления начинает падать. Большая употребительность тематических слов в газетных текстах понятна. Но высокая частотность общих слов, т.е. тех, которые являются нейтральными и характерными для любого стиля, ведет к появлению бюрокронимов.

Наиболее употребительными в русских текстах являются предлоги, союзы, местоимения, а также тематические слова: *управление, республика, российский, чувашский, президиум, федеральный, год*, являющиеся типичными для газетной лексики. Среди стилистически нейтральных слов высокую ступень сатиации занимает глагол *быть*. Данному слову предполагается реальная угроза подвергнуться семантическому износу.

В чувашских газетах высокочастотными являются знаменательные части речи. Многократно повторяющимися явились такие тематические слова, как *чаваш(чувашский), республика, музей, Рассей. (Россия), сул (год), патшалаш (государство), канаш (совет), курав (выставка), саккун (закон), президент, халах (народ), хусалах (хозяйство), пёрлешу (союз), зал, ёсчен (рабочий), хула (город), ял (деревня), семинар, фонд, вёрепту(образование), депутат, сёр (земля), завод, район*. Стилистически немаркированные слова из-за высокой частотности стали штампами: *ёс (работа), пул (состояться), унта (там), хай(сам), ёсле (работать), ирт (провести), ку (этот), сав (тот), ус (открывать), чи (самый), сёне (новый), пар (отдавать), пыр (приходить), малтан (сначала), кала (говорить), юлашки (последний), паллаш (знакомиться), ту (делать), пирён (у нас), хатёрле (готовить), пысак (большой)*. Последнее слово употребляется в чувашских текстах для выражения признака самых различных понятий: *"пысак ёс" (большая работа), "пысак зал" (большой зал), "пысак пулам" (знаменательное событие), "пысак уяв" (большой праздник)*. Полученные результаты позволяют констатировать, что в чувашских газетах значительно слабее используются синонимические возможности языка.

В заключении сформулированы общие выводы, полученные в результате сопоставительного лингвостатистического анализа языковых единиц региональной русской и чувашской публицистики. В диссертации утверждается: в обоих языках имеются специфические средства создания газетной речи, обнаруживаются стилистические, жанровые особенности сравниваемых языков, а также их типологическое своеобразие. В чувашской публицистике границы между жанрами более размыты и имеют менее строгую дифференциацию по сравнению с русской. Отмечается, что квантитативный подход создает перспективу превращения обычной структурной модели языка в структурно-вероятностную, отличающуюся большей реалистичностью, а также способствует решению актуальной для всех языков проблемы создания мета-языка.

Перспективным направлением является изучение чувашского языка в сопоставлении с другими, в историческом прошлом имевшими с ним тесные контакты или даже являвшимися родственными, но в настоящее время территориально разделенными. Это преимущественно восточные языки — тунгусо-маньчжурские, корейский, китайский, японский. Вместе с тем актуальным представляется сопоставление чувашского языка с пребывавшими с ним в контакте до последнего времени коми, марийским, мордовским, удмуртским и др. Необходимо изучение общих и отличительных параметров с новоконтактирующими международными языками в целях лучшего осмысления особенностей родного языка и оптимизации преподавания лингвистических дисциплин. И, наконец, благодаря такого рода лингвостатистическим исследованиям появляется надежда, что в научный обиход будет вовлечен статистический материал по чувашскому языку, до последнего времени не замеченный в общетеоретических квантитативных исследованиях.

Основное содержание диссертации отражено в следующих публикациях:

1. Предупреждение о нарушении языковой экологии // Болгаро-чувашская цивилизация: прошлое, настоящее, будущее: Сб. ст.- Чебоксары, 2000.-С. 138-142.

2. Вероятность в лингвистике // Труды молодых ученых и специалистов, Чебоксары, 2001.-С. 49.- 52.

3. Категории количества и качества в языкознании // Экономическая, политическая и духовная трансформация Приволжского федерального округа: Сб. статей Поволж. науч. конф.-Чебоксары, 2001.-С. 114-122.

4. Корреляционный анализ частей речи в русских и чувашских газетных текстах // Актуальные проблемы чувашского языка: Материалы конф., посвящ. 60-летию со дня рождения д-ра филол. наук проф. В.И. Сергеева.- Чебоксары, 2002.-С. 121-124.

5. Специфика лингвистических объектов на пути формализации языка: Материалы конф., посвящ. 30-летию со дня образования Чуваш. гос. ун-та им. И.Н. Ульянова.- Чебоксары, 2002.- С. 98 - 102.